

Received January 27, 2022, accepted February 8, 2022, date of publication February 15, 2022, date of current version March 3, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3151789

# Respiratory Sound Classification: From Fluid-Solid Coupling Analysis to Feature-Band Attention

FUCHUAN TONG<sup>1</sup>, LINGLING LIU<sup>1</sup>, XINGJIA XIE<sup>1</sup>, QINGYANG HONG<sup>2</sup>, (Member, IEEE), AND LIN LI<sup>1</sup>, (Member, IEEE)

<sup>1</sup>School of Electronic Science and Engineering, Xiamen University, Xiamen 361005, China

<sup>2</sup>School of Informatics, Xiamen University, Xiamen 361005, China

Corresponding author: Lin Li (lilin@xmu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61876160 and Grant 62001405.

**ABSTRACT** Based on respiratory sound production mechanisms, we study the relationship between airflow characteristics in the bronchi and sound pressure spectrum curves to implement an end-to-end respiratory sound classification system with a feature-band attention module. First, we analyse fluid-solid coupling simulations of the bronchi and execute acoustic simulations to obtain the spectrum curves of the bronchi at the sound pressure level. Then, based on the spectrum characteristics of the bronchi, we propose an attention strategy to refine the acoustic features with adaptive weights. In addition, we introduce a feature-band attention module to ResNet-based networks with a squeeze-and-excitation block. Finally, we perform experiments on the ICBHI public database to classify respiratory sounds into one of four classes: normal, wheezes, crackles, and both (wheezes and crackles). The results show that our proposed system exhibits superior performance compared with the baseline system. This type of feature learning strategy is useful for exploring the distinct characteristics of different types of respiratory sounds.

**INDEX TERMS** Fluid-solid coupling, attention learning, end-to-end system, respiratory sound classification, squeeze-and-excitation.

## I. INTRODUCTION

In recent years, chronic respiratory diseases have spread all over the world with high prevalence rates and recurrent attacks. Asthma and chronic obstructive pulmonary disease (COPD) are the main representative diseases [1]. Patients with asthma and COPD require long-term treatment and daily monitoring, which is challenging due to limited medical resources, as clinicians must auscultate for respiratory sounds during each visit at internal medicine departments. Fortunately, with the rapid development of machine learning technology, many automatic respiratory sound classification systems have been launched to pave a new way forward in aiding clinical diagnoses and treatment of respiratory diseases [2].

The commonly used framework for respiratory sound classification includes two parts: feature extraction for respiratory sound signals and a classification model. For feature

extraction, many methods have been adopted for extracting acoustic features, including entropy-based features [3], Short-Time Fourier Transform (STFT) [4], spectrogram [5], Mel-Filter banks (FBank) [6], Wavelet analysis [7], Perceptual Linear Prediction (PLP) [8], and Mel-Frequency Cepstral Coefficient (MFCC) [9]. For the classification model, some discriminative models and generative models have been employed, such as the k-Nearest Neighbour (k-NN), Gaussian Mixture Model (GMM), Hidden Markov Model (HMM), and Support Vector Machine (SVM). As early as 1997, some experts began using autoregressive models for extracting spectrum features from respiratory sounds, and they established multiple k-NN classifiers to recognize respiratory sounds [10]. Improved performance was achieved by HMM-based and SVM-based systems [11]–[13].

Recently, deep neural networks have achieved promising performance, and continuous optimizations have been made in algorithms and applications. In our previous work [14], we combined HMM and deep neural network architecture to build a classification model for respiratory sounds.

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino<sup>1</sup>.

Some studies have focused on transfer learning strategies by utilizing the pretrained VGG16 model on image datasets [15]–[17]. More complex CNN-based systems with a model tuning strategy and a data augmentation technique were proposed to obtain better performance in the respiratory sound classification task [18]–[20]. However, the abovementioned systems directly adopt conventionally used acoustic parameters in the speech classification field or spectrogram representation for image classification. To improve the performance of respiratory sound classification systems, we aim to explore the innate characteristics of respiratory sounds by physically modelling the bronchi.

In the geometric modelling of bronchi, scholars started with the bronchi's geometric characteristics and modelled the bronchi through a variety of methods. For example, they studied airflow movement and particle deposition in the trachea. In 1982, Haselton and colleagues simulated the airflow state in the windpipe by using a symmetrical bifurcation bronchi model [21]. In 1989, based on the bifurcated pipe model, Snyder and Olson focused on the flow velocity distribution of the airflow near each bifurcation point. The research results show that strong shear stress may directly cause the airflow to produce secondary flow vortices [22]. Cheng *et al.* simulated particle deposition through oral and bronchi models [23].

After conducting geometric modelling of the bronchi, the airflow simulation process requires the related theory of fluid mechanics. In 1755, the Euler equation [24] was proposed to describe fluid motion under ideal conditions. Because the Euler equation assumes that fluid is a non-viscous ideal object, its applications in practical engineering are limited. In the 19th century, Navier and Stokes considered the viscous fluid movement in the boundary layer and defined the Navier-Stokes equation, also referred to as the N-S equation [25], which is the theoretical cornerstone of modern fluid mechanics. However, the N-S equation solution is also a challenge. Experts recommend turbulence simulation methods, including direct numerical simulation and indirect numerical simulation. Examples of indirect simulation methods are the Large Eddy Simulation (LES) and Reynolds averaged Navier-Stokes (RANS). The  $k - \omega$  model and the  $k - \varepsilon$  model are two mainstream modelling solutions in the RANS method, in which  $k$  represents turbulent kinetic energy,  $\omega$  is the frequency of turbulent decay processes, and  $\varepsilon$  denotes the turbulent energy dissipation rate. For the law of airflow movement in the bronchi, the above numerical simulation methods are usually used. For example, Zhao and colleagues used the  $k - \omega$  model to analyse airflow characteristics in the upper respiratory tract of the human body [26]. Mihai Mihaescu *et al.* simulated the established human airway model and found that the RANS method was not suitable for predicting the anisotropic fluid movement. For the airflow simulation of the human bronchial airway, the LES large eddy simulation method captures microscopic characteristics of airflow and fluid movement [27], [28].

However, most of the above numerical simulation studies on bronchial airflow patterns ignored the influence of bronchial tube deformation and fluid-solid coupling and directly set the bronchial wall structure as a rigid wall. In actual situations, even small movements of the bronchial wall have a great impact on airflow in the bronchial tubes. Therefore, it is beneficial to consider the influence of the wall on bronchial airflow under actual physiological conditions within the human body. To complement this research gap, we modelled the bronchi with fluid-solid coupling and simulated the airflow in the bronchi, which provides basal data for subsequent acoustic modelling.

The acoustic characteristics of respiratory sound signals have a strong correlation with airflow characteristics in the lungs. However, most of the early studies on acoustic modelling of bronchi only analysed the relationship between respiratory sound production mechanisms and bronchial airflow but failed to map relationships between respiratory sound signal pressure levels and bronchial airflow. For example, Forgacs analysed the sound generation mechanisms of different types of respiratory sounds in the lungs, classified them based on corresponding airflow patterns in the tube, and found that there are three modes of bronchial airflow: laminar, turbulent, and vortex. The research results of Hardin *et al.* on vortex vocalisation in the respiratory system further demonstrate Forgacs' vortex theory, which illustrates that the vortex phenomenon is generated when the airflows from the small bronchi to the large bronchi [29]. The respiratory sound source is broadband noise, so the vortex is the main source of respiratory sounds.

In response to the problem of continuous wheezing, Vaz and Thakor proposed the airway tremor theory, which reveals the relationship between wheezing and airflow [30]. Xu *et al.* used the finite difference method to study the production model and further clarified the theoretical production mechanism of wheezes; wheezes are produced by the interaction between the airway wall and airflow during the movement of respiratory airflow [31]. In 2017, Messner *et al.* used linear predictive cepstral coefficients and polynomial regression to map the relationship between intrabronchial airflow and generated respiratory sounds [32], and they detected the respiratory phase of respiratory sound signals. It is assumed that the airway pathophysiology can be detected to diagnose respiratory diseases [33].

In the field of aeroacoustics, the Ffowcs Williams-Hawkings (FW-H) equation is usually used to determine the relationship between the sound pressure of sound signals and airflow movement [34]. Zhang *et al.* combined the LES and FW-H equations to predict the noise spectrum of the wings of an aircraft [35]. After using different turbulence calculation methods to study the influence of turbulence on an aircraft's wings, Li *et al.* used the FW-H method to investigate the influence of the thickness of the wings' trailing edges on noise characteristics [36]. These studies in the field of aeroacoustics inspired us to map the relationships between

bronchial structures and respiratory sound acoustic characteristics. Without considering the bronchial wall, we mapped the relationships between bronchial structures and respiratory sounds and achieved some conclusions [37]. We plan to further explore the aeroacoustic phenomenon with the effect of bronchial walls.

In this paper, we study respiratory sound production mechanisms based on physical models of bronchi and airflow movement. We simulated airflow characteristics in the bronchi and the sound pressure spectrum. Based on the relationships between acoustic spectra and lung pathology, we further developed an end-to-end respiratory sound classification system with a feature-band attention module. This paper is organized as follows. Section II describes the related formulas for bronchial and acoustic modelling and the baseline system in this study. Section III describes the methods proposed in this study, including bronchial physical modelling and optimized automatic classification systems. Section IV illustrates the experimental settings and analyses the simulation results of physical and acoustic modelling. Section V discusses the performance of respiratory sound classification systems. Finally, Section VI concludes the paper.

## II. RELATED WORKS

To better understand the mathematical content in this paper, we define consistent notations. The subscript  $n$  denotes the direction outside the wall of the tube. The subscripts  $i, j$ , and  $k$  denote the directions of fully Cartesian coordinates. The vectors with the superscript  $b$  represent the parameters of the bronchial tube wall.

### A. FLUID CONTROL EQUATION

Airflow in the bronchi is in a roughly viscous flow state, which is isothermal and incompressible. The continuity equations for respiratory airflow movement in the tube and the N-S control equation are described in [38] and formulated as follows:

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (1)$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial (u_i u_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} + h, \quad (2)$$

where  $u_i$  and  $u_j$  are the velocities in directions  $i$  and  $j$ , respectively;  $x_i$  and  $x_j$  are the displacements in directions  $i$  and  $j$ , respectively.  $\rho$  is the airflow density,  $\nu$  is the kinematic viscosity coefficient of airflow in bronchi, and they are related to viscosity  $\mu$ ,  $\mu = \rho \times \nu$ . The viscosity of airflow  $\mu$  is set to  $1.89 \times 10^{-5}$  Pa · S.  $p$  is the pressure value per unit area produced by the flow field in the bronchi.  $t$  indicates time,  $h$  represents volume force,  $h = \rho g$ , and  $g = 9.8$  N/kg.

In the flow field simulation in the bronchi, the standard  $k - \varepsilon$  model is adopted for steady simulation, and the LES simulation is employed to solve the N-S equation. In the LES simulation, the filtered continuity equation for airflow movement in the bronchi and the N-S control equation are

written as:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0, \quad (3)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial (\bar{u}_i \bar{u}_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j}, \quad (4)$$

where  $\bar{u}_i$  denotes the mean of  $u_i$ , and  $\bar{u}_i \bar{u}_j$  is written as  $\bar{u}_i \bar{u}_j = \bar{u}_i \bar{u}_j + (\bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j)$ . We concatenate (3) and (4):

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial (\bar{u}_i \bar{u}_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} + \frac{\partial}{\partial x_j} (\bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j). \quad (5)$$

### B. SOLID CONTROL EQUATION

We also consider the influence of bronchial wall deformations on bronchi airflow. According to nonlinear continuum mechanics, the control equation for the solid bronchial wall is as shown in [38]:

$$\rho^b \frac{du^b}{dt} = \nabla (\mathbf{F} \cdot \mathbf{Y}^b) + \rho^b f^b, \quad (6)$$

where  $\rho^b$  is the bronchial wall density,  $u^b$  is the motion velocity of the bronchial wall,  $\mathbf{Y}^b$  is the stress tensor,  $f^b$  denotes stress near the boundary of the bronchial tube wall, and  $\mathbf{F}$  is the deformation gradient of the bronchial wall, which is related to wall displacement.

### C. SPECTRUM SIMULATION OF SOUND PRESSURE LEVELS

We use the FW-H equation to simulate the sound pressure level spectrum. The FW-H equation is as follows:

$$\frac{1}{a_0^2} \cdot \frac{\partial^2 p'}{\partial t^2} - \nabla^2 p' = \frac{\partial}{\partial t} (A \delta (f)) - \frac{\partial}{\partial x_i} (F_i \delta (f)) + \frac{\partial^2}{\partial x_i \partial x_j} (T_{ij} H (f)), \quad (7)$$

where  $a_0$  represents the sound velocity in the bronchi,  $\delta (f)$  represents the Dirichlet function,  $H (f)$  is the step function, and  $p'$  represents the sound pressure at an observation point in the bronchi, which can be arbitrarily designated.

The right side of (7) represents the source term generated by bronchial aerodynamic noise, and the monopole source term is expressed as:

$$A = \rho_0 v_n + \rho (u_n - v_n), \quad (8)$$

where  $\rho_0$  is the airflow density value without perturbation. Since airflow in the bronchi can be regarded as incompressible viscous flow, in this paper,  $\rho = \rho_0$ .

The dipole source term is expressed as:

$$F_i = Y_{ij} n_j + \rho u_i (u_n - v_n), \quad (9)$$

where  $n_j$  represents the unit normal vector pointing to the outer area of the flow field surface. The fluid velocity component perpendicular to the integration surface is represented by  $u_n$ , and  $v_n$  represents the velocity component of integration

surface movement.  $Y_{ij}$  represents the stress tensor of the fluid domain:

$$Y_{ij} = \bar{p}\delta_{ij} - \mu \left[ \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} - \frac{2}{3} \frac{\partial \bar{u}_k}{\partial x_k} \delta_{ij} \right], \quad (10)$$

where  $\mu$  represents the viscosity of the airflow in the lungs, and  $\bar{p}$  is the average sound pressure level.

The quadrupole source term  $T_{ij}$  is expressed as:

$$T_{ij} = Y_{ij} + \rho \bar{u}_i \bar{u}_j - a_0^2 (\rho - \rho_0) \delta_{ij} = Y_{ij} + \rho \bar{u}_i \bar{u}_j. \quad (11)$$

Based on equation (7), the sound pressure  $p'$  of a designated observation point of the bronchi is calculated. Then, the effective sound pressure  $p(e)$  is obtained [39]:

$$p(e) = \sqrt{\frac{1}{T} \int_0^T (p')^2 \cdot dt}. \quad (12)$$

We calculate the sound pressure level  $SPL$  for a designated observation point [39]:

$$SPL = 20 \lg \frac{p(e)}{p(ref)}, \quad (13)$$

where  $p(ref)$  is the reference sound pressure and generally  $p(ref) = 2 \times 10^{-5}$  Pa.

### D. DEEP NEURAL NETWORK ARCHITECTURE

We take ResNet [40] as the backbone network for the baseline respiratory sound classification system. The basic component of the ResNet network is a residual module, which uses a jump-connect structure (also known as identity mapping). Since the ResNet network is composed of many residual modules stacked together, it is easy to modify and expand the network structure.

## III. PROPOSED METHODS

To explore the relationship between sound generation on the bronchial structure and the feature representations of respiratory sound signals, we propose an end-to-end respiratory sound classification framework with a feature-band attention module, which is obtained from the fluid-solid coupling simulation of a bronchial model.

### A. BRONCHIAL MODELLING

#### 1) FLUID-SOLID COUPLING BOUNDARY CONDITIONS FOR BRONCHI

Fluid-solid coupling modelling of bronchi requires boundary conditions: kinematic conditions, fluid velocity conditions, and dynamic conditions. Specifically, on the fluid-solid coupling interface between the bronchial fluid domain and the solid domain, the following conditions (a)-(c) should be satisfied [38]:

##### a: KINEMATICS CONDITION

The corresponding mass point between the fluid boundary domain and the solid domain (i.e. the wall of the bronchi) shares a consistent displacement. Namely,  $x^b|_0 = x|_0$ ,

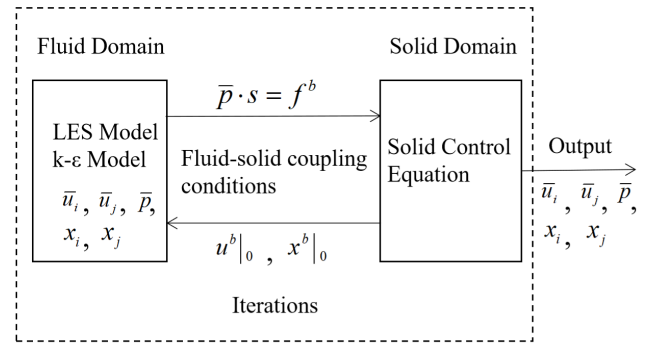


FIGURE 1. The iteration process for fluid-solid coupling.

where  $x^b|_0$  represents the boundary displacement of the bronchi's solid domain close to the fluid boundary surface, and  $x|_0$  represents the boundary displacement of the fluid domain close to the solid boundary surface.

##### b: FLUID VELOCITY CONDITION

The fluid-solid coupling interface adopts a non-slip wall boundary condition, and the boundary velocity of the bronchial fluid domain and the boundary velocity of the bronchial wall should satisfy this condition. Namely,  $u^b|_0 = u|_0$ , where  $u^b|_0$  represents the boundary velocity of the bronchial wall close to the fluid boundary surface, and  $u|_0$  represents the boundary velocity of the fluid domain close to the boundary surface of the bronchial wall.

##### c: DYNAMIC CONDITION

The balanced force condition is  $f^b \cdot n^b = f \cdot n$ , where  $f^b$  represents the stress of the solid domain close to the fluid boundary surface, and  $f$  represents the stress of the fluid domain close to the solid boundary surface.  $n^b$  is the normal vector of the bronchial wall, and  $n$  is the normal vector of the fluid boundary in the bronchi.

The iterative calculation flowchart for the fluid-solid coupling of the bronchial model is shown in Fig. 1. The standard  $k-\epsilon$  model and the solid control equation are used to simulate the flow field in the bronchi, to calculate the flow field velocity  $\bar{u}_i$  and  $\bar{u}_j$ , to determine flow field displacement  $\bar{x}_i$  and  $\bar{x}_j$ , and to find the pressure  $\bar{p}$  per unit area of the flow field. When the flow field reaches a steady flow state, we select LES to simulate the flow field more precisely.

#### 2) SOUND PRESSURE SPECTRUM CURVE FOR RESPIRATORY SIGNALS

To solve the FW-H equation, we substitute the pressure  $\bar{p}$ , the flow field velocity  $\bar{u}_i$  and  $\bar{u}_j$ , and the flow field displacement  $x_i, x_j$  into equation (10) to obtain  $Y_{ij}$ . Then, we substitute  $Y_{ij}$  into (9) and (11) to determine  $F_i$  and  $T_{ij}$ , and we substitute  $F_i, Y_{ij}$ , and  $T_{ij}$  into (7) to find  $p'$ . The effective sound pressure  $p(e)$  is obtained after the root mean square calculation of the instantaneous sound pressure  $p'$  at the designated observation

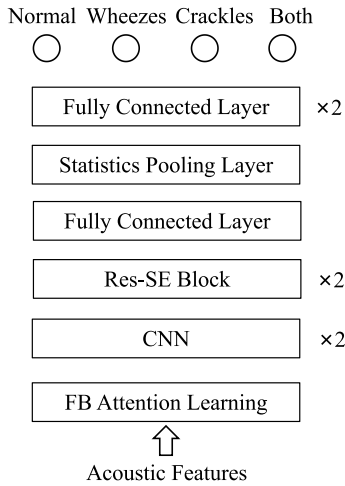


FIGURE 2. The overall framework.

point over time, and then the sound pressure level  $SPL$  at the designated observation point is calculated in (13).

### B. RESPIRATORY SOUND CLASSIFICATION FRAMEWORK

#### 1) SYSTEM STRUCTURE

We design end-to-end respiratory sound classification systems based on the ResNet network. In contrast to the existing systems, we do not copy the acoustic features from the speech domain or pretrained neural networks for image recognition. We introduce the feature acoustic attention mechanism to the ResNet-backbone structure to strengthen the representation of frequency band characteristics among different respiratory sound signals. Due to the feature attention module that is inferred by bronchial modelling and simulation, the spectrum characteristics of the respiratory sound signals are adaptively weighted to further improve the classification performance. The overall framework is shown in Fig. 2.

#### 2) CHANNEL-WISE ATTENTION MODULE

The whole structure of the ResNet baseline includes two 1-dimensional convolutional layers (Conv1D), two ResNet blocks that operate on the frame level, a statistic pooling layer that calculates the mean and standard deviation of each sample along the time-axis, some fully-connected (FC) layers, and an output layer with four nodes for four classes. Based on the ResNet baseline, we adopt the Squeeze-and-Excitation (SE) module [41] to capture the dependency between different feature channels and adjust the weight of each feature channel adaptively. The output of the residual block flows through the SE block before the skip connection, where  $T$  denotes the sequence length,  $F$  is the feature dimension,  $C$  denotes the number of channels of the residual block, and parameter  $r$  is a reduction ratio for controlling the computational cost of the SE block.

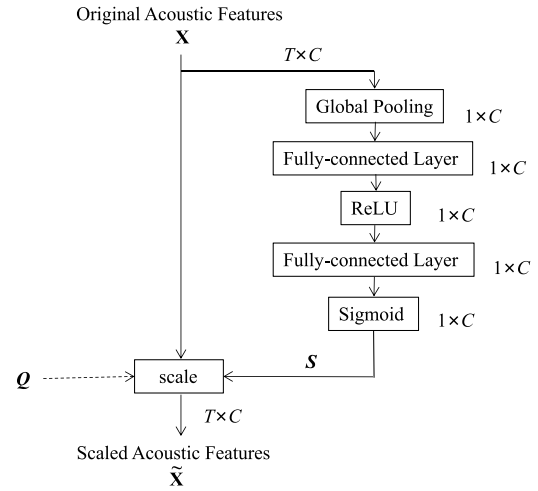


FIGURE 3. Band attention module.

#### 3) FEATURE-BAND ATTENTION LEARNING

Based on the analysis of spectrum curves for respiratory sound pressure levels, we propose two types of feature-band attention modules to obtain important frequency band information in respiratory sound signals, namely, a feature-band attention module (FB attention module) and a feature-band with a Q-parameter band attention module (FBQ attention module). With the FB attention module, the weight distribution of frequency bands in respiratory sound signals is obtained by the network’s training process, and the scale operator weights the acoustic feature vectors. In the FBQ attention module, the spectrum characteristics for respiratory sound pressure simulation analyses are considered to include a  $Q$  vector. Then, the scale operator conducts band-wise multiplication with the  $Q$  vector to reinforce informative features and suppress less useful features. The structures of these two feature-band attention module types are shown in Fig. 3, where  $X$  represents the original acoustic features of respiratory sound signals, and  $\tilde{X}$  represents scaled acoustic features with a feature-band attention module.

Given  $C$  bands of Mel-filter banks and  $T$  frames for each sample, the acoustic feature vectors in each Mel-filter band are averaged through the global pooling layer:

$$z_c = \frac{1}{T} \sum_{t=1}^T x_{t,c}, \quad (14)$$

where  $x_{t,c}$  represents the value of the  $c$ -th filter band in the  $t$ -th frame from the original acoustic features.  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t, \dots, \mathbf{x}_T]$ ,  $\mathbf{x}_t = [x_{t,1}, \dots, x_{t,c}, \dots, x_{t,C}]$ .  $z_c$  represents the mean of the  $c$ -th filter band, and  $\mathbf{Z} = [z_1, z_2, \dots, z_c, \dots, z_C]$ . The band-wise weight  $\mathbf{S} = [s_1, s_2, \dots, s_c, \dots, s_C]$  is calculated through a series of non-linear transformations.

$$\mathbf{S} = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{Z})), \quad (15)$$

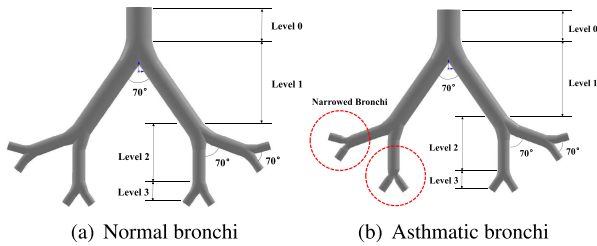


FIGURE 4. The bronchi geometric model.

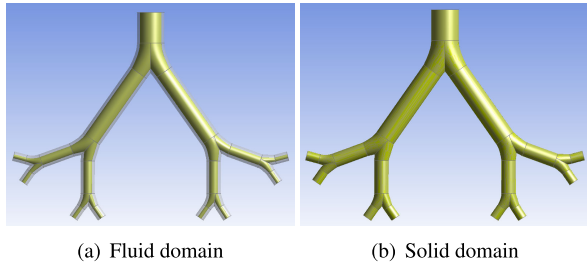


FIGURE 5. The fluid domain and solid domain of the bronchi geometric model.

where  $\sigma(\cdot)$  represents the sigmoid function,  $\delta(\cdot)$  is the ReLU function, and  $\mathbf{W}_1$  and  $\mathbf{W}_2$  are the network weights of the two fully-connected layers.

For the FB attention module, the weight  $S$  dots the original acoustic features  $\mathbf{X}$  to obtain new scaled acoustic features, and  $\tilde{\mathbf{X}} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_t, \dots, \tilde{x}_T]$ ,  $\tilde{x}_t = [\tilde{x}_{t,1}, \dots, \tilde{x}_{t,c}, \dots, \tilde{x}_{t,C}]$ :

$$\tilde{x}_{t,c} = F_{\text{scale}}(x_{t,c}, s_c) = s_c \cdot x_{t,c}. \quad (16)$$

In the FBQ attention module, the parameter  $\mathbf{Q} = [q_1, q_2, \dots, q_c, \dots, q_C]$  is used to control the weight  $S$ , and then dots the original acoustic features  $\mathbf{X}$  to obtain a new scaled acoustic feature  $\tilde{\mathbf{X}}$ :

$$\tilde{x}_{t,c} = F_{\text{scale}}(x_{t,c}, s_c, q_c) = s_c \cdot x_{t,c} \cdot q_c. \quad (17)$$

We introduce these two types of feature-band attention modules to the SE-ResNet system and design the FB-SE-ResNet system and the FBQ-SE-ResNet system.

#### IV. BRONCHIAL PHYSICAL MODELLING AND SIMULATION RESULTS

##### A. EXPERIMENTAL SETUP

###### 1) BRONCHIAL MODELLING

Based on the airway tree model proposed by Weibel, we rendered the geometric structure of normal and asthmatic bronchi with *SolidWorks* modelling software. Since the vortex of the fluid movement in the large bronchi greatly contributes to creating respiration sounds in the bronchi, we select 0-3 level bronchi in the Weibel model for geometric modelling, and the wall thickness is 1.65 mm [42]. The specific parameters of the bronchial geometric model are shown in Table 1.

TABLE 1. Specific parameters of bronchial geometric model (unit:mm).

Level	Diameter	Length
0	18.0	20.0
1	12.2	47.6
2	8.3	19.0
3	5.6	7.6

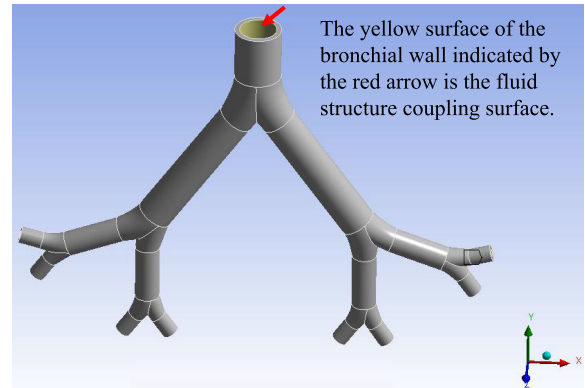


FIGURE 6. The fluid-solid coupling surface in the solid domain.

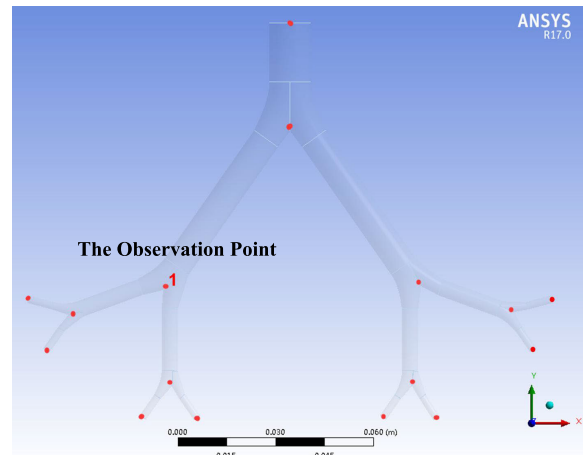
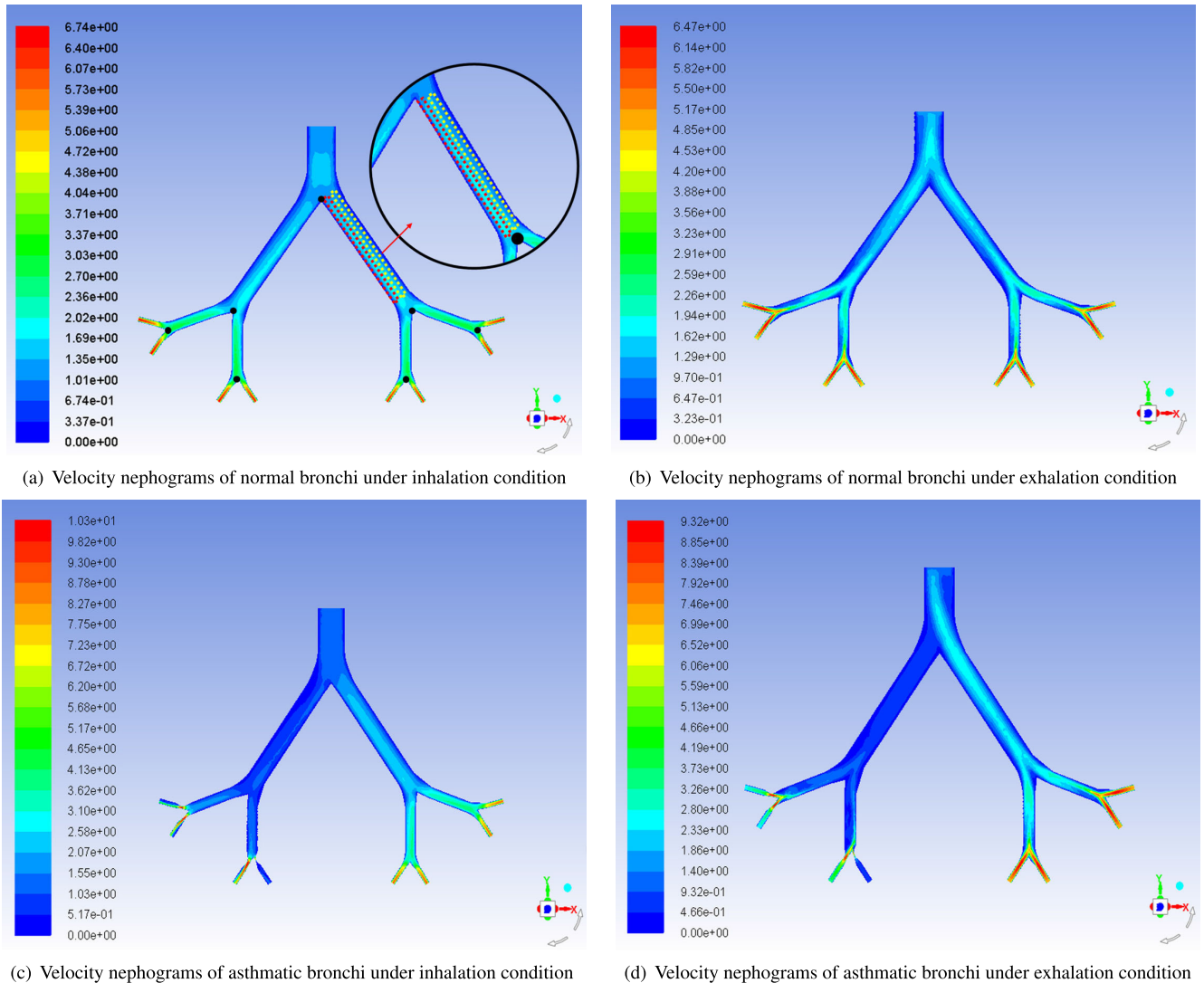


FIGURE 7. The observation point of sound pressure levels on the bronchi.

Fig. 4 (a) shows the normal bronchi geometric model. We also narrowed some specific bronchi to obtain the asthmatic bronchi structure, as shown in Fig. 4 (b), and it is emphasized by two red circles. The fluid domain and the solid domain of the bronchi are shown in Fig. 5, where the solid domain of the bronchi in Fig. 5 (b) is the shaded part outside the yellow area in Fig. 5 (a).

###### 2) FLUID-SOLID COUPLING MODELLING

We used *ANSYS* software to simulate the fluid-solid coupling model. First, we executed the Fluid Flow (Fluent) workbench in *ANSYS* software to simulate the airflow velocity and pressure of the bronchi fluid domain, and we utilized the Transient Structure for the physical qualities of the bronchi's solid domain, including the solid wall displacement of the bronchi



**FIGURE 8.** Velocity nephograms of normal bronchi and asthmatic bronchi (velocity unit: m/s).

and the velocity of the tracheal wall. Then, the System Coupling module in ANSYS is used to transfer data between the fluid domain and the solid domain.

In the Fluent workbench, we set the inlet boundary condition of the bronchi so that airflow enters the bronchial inlet at a certain speed with the software option *Velocity Inlet*. The air inlet velocity was 1.2 m/s, and it was evenly distributed on the inlet surface with an airflow turbulence rate that is set to 10%. The outlet boundary conditions of the eight outlets for the bronchi are all set to *Pressure Out* (software option), and the relative pressure of the outlet was 0 [28]. We assumed that this fluid-structure interaction simulation experiment was carried out under the physiological condition of a human body temperature of 37 °C. The airflow density  $\rho$  is set to 1.1 kg/m<sup>3</sup>, and the airflow viscosity  $\mu$  is set to 1.89 × 10<sup>-5</sup> Pa · S. We used the standard  $k - \varepsilon$  model to calculate the airflow velocity until the bronchial fluid domain

reached a steady flow state, and then we adopted the LES simulation to further calculate the flow field information with higher simulation accuracy.

In the Transient Structural module of ANSYS, we simulated the pressure and displacement of the bronchial wall during respiration. In the experiments, the density  $\rho^b$  of the bronchial wall was set to 1,060 kg/m<sup>3</sup>, Poisson's ratio  $\sigma$  was 0.4, and the elastic modulus was set to 0.9 MPa, which follows Hooke's law [43]. In the solid domain, the inner wall of the bronchi is set as a fluid-solid coupling surface. As shown in Fig. 6, the fluid-solid coupling surface is marked in yellow and indicated by a red arrow.

Finally, we utilized the System Coupling module in ANSYS software and linked the simulation data to the Fluid Flow module for the fluid domain and the Transient Structural module for the solid domain. To prevent extreme deformations in the grid or excessive distortion rates in the simulation

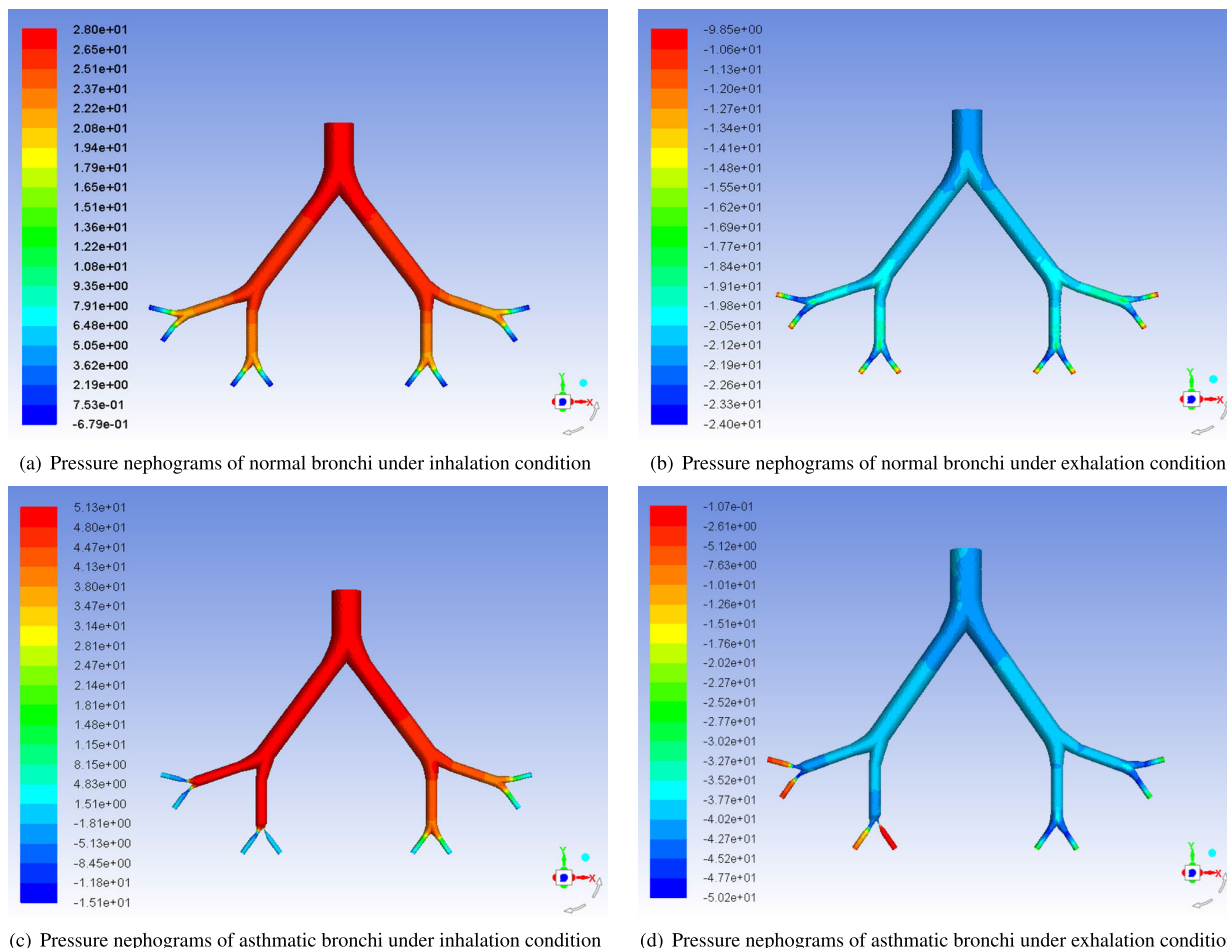


FIGURE 9. Pressure nephograms of normal bronchi and asthmatic bronchi (pressure unit: Pa).

process, we employed a smoothing method to control the grid of the fluid-solid coupling surface and to correct the deformed grid in real time.

### 3) ACOUSTIC SIMULATION SETTINGS

In the acoustic module, the acoustic field information is calculated according to the airflow velocity  $\bar{u}_i$  and  $\bar{u}_j$  of the bronchi fluid domain at each time step and the pressure intensity  $\bar{p}$  of the flow field.

In clinical auscultation, the position where the auscultation head is placed roughly corresponds to the level 2 bronchi in the Weibel model. We chose position “1”, as shown in Fig. 7, as the observation point for the sound pressure level. The time step is set to  $2.5 \times 10^{-4}$  s in the simulation of sound pressure levels. When the simulation converged, the sound field simulation was completed. Then, we used the Fourier Transform module to obtain the sound pressure level spectrum curve.

## B. SIMULATION RESULTS

### 1) AIRFLOW VELOCITY OF BRONCHI

Fig. 8 shows velocity nephograms of normal bronchi and asthmatic bronchi under inhalation and exhalation conditions, which were scanned in the plane on the central axis of

the bronchi. As shown in Fig. 4, we segmented the bronchi into four regions, which were referred to as levels 0 through 3, and we found that airflow velocity increases as it moves through these ascending levels.

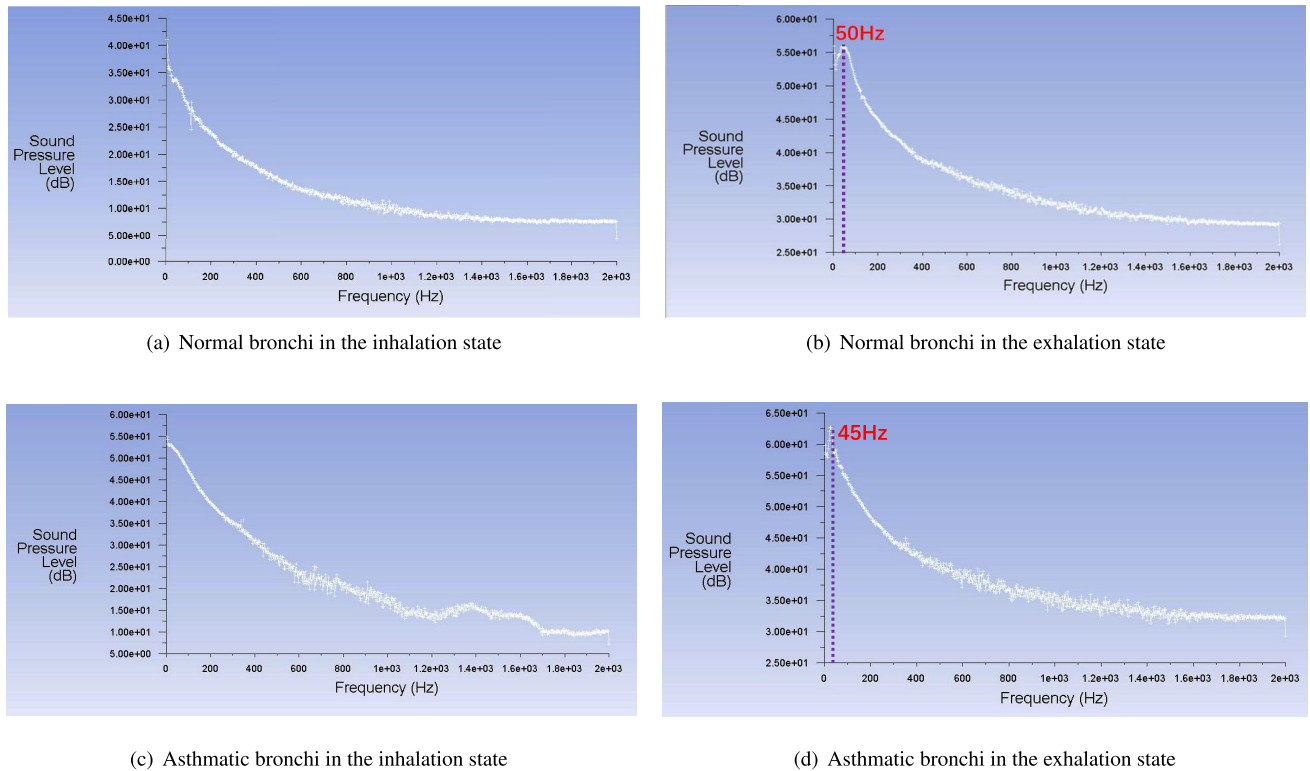
For normal bronchi, a flow rate imbalance occurs, wherein the airflow velocity on the side close to the bifurcation point is faster, and the airflow velocity at the central axis of the bronchi is higher than that on both sides.

For asthmatic bronchi, the overall velocity of airflow on the blocked side of the bronchus is greater than that on the unblocked side of the bronchus. For example, the maximum flow velocities of the blocked bronchus during inhalation and expiration are approximately 10.3 m/s and 9.32 m/s, respectively, whereas the maximum values of the normal bronchus flow velocity in the inhalation state and the expiration state are 6.74 m/s and 6.47 m/s, respectively.

### 2) BRONCHIAL WALL PRESSURE

We compared the pressure nephograms of normal bronchi and asthmatic bronchi under inhalation and exhalation conditions, which were scanned in the plane on the central axis of the bronchi. As shown in Fig. 9, in the inhalation state, the maximum bronchial wall pressure is in the entrance region,





**FIGURE 10.** The spectrum curves of the sound pressure levels of normal bronchi and asthmatic bronchi.

and the wall pressure decreases as the airflow moves through levels 0-3 of the segmented bronchi model. However, in the exhalation state, the maximum wall pressure of the bronchi is at the outlet region, and the wall pressure increases as the airflow moves through levels 0-3 of the segmented bronchi model.

Furthermore, we found that the wall pressure of asthmatic bronchi is more uneven than that of normal bronchi. The overall pressure range on the blocked side of the bronchus is greater than that on the unblocked side of the bronchus (as well as that of the normal bronchi), which is consistent with the phenomenon that patients with physiological asthma breathe more difficultly.

### 3) SOUND PRESSURE LEVEL SPECTRUM CURVE

In the acoustic simulation process, we compared the sound pressure level spectrum curves of normal bronchi and asthmatic bronchi in both inhalation and expiration states, which were obtained from observation point 1. Fig. 10 shows that spectrum curves of normal respiratory sounds and asthmatic sounds are mainly distributed below 2,000 Hz, but sound pressure levels of asthmatic bronchi are higher than those of normal bronchi both in the exhalation and inhalation state, as well as the distribution differences in spectrum peaks. For example, the spectrum curves reach maximum peaks at 45 Hz (asthmatic bronchi) and 50 Hz (normal bronchi) under exhalation conditions. Under inhalation conditions, the sound pressure level spectrum curve of normal bronchi ranges

from 0 to 1,000 Hz and thereafter declines until stabilizing at approximately 1,000 Hz. However, the sound pressure level spectrum curve of asthmatic bronchi still fluctuates with low pressure level values during the frequency range of 1,000 Hz to 2,000 Hz. These spectrum peak distributions are nearly identical to clinical statistics, which benefits the design of the feature-band attention module proposed in Section III.

## V. CLASSIFICATION RESULTS AND ANALYSIS

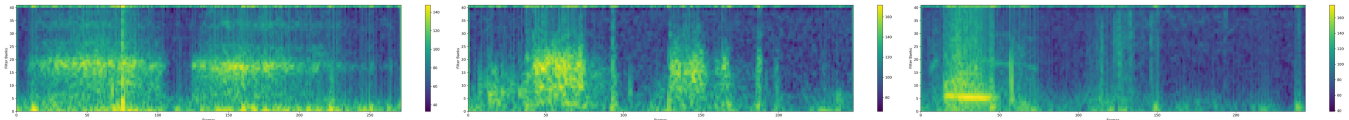
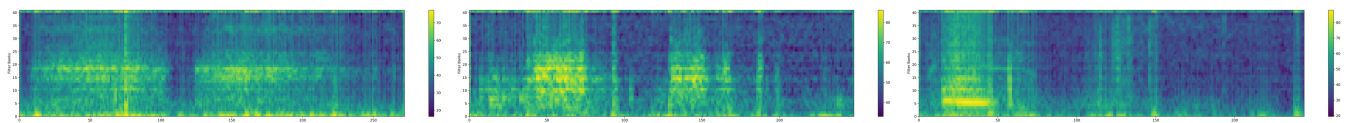
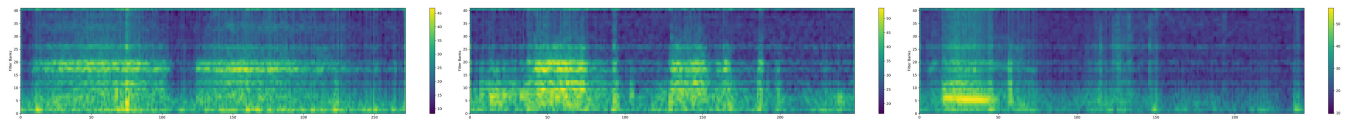
### A. DATASET

To verify the proposed methods, we used the International Conference on Biomedical and Health Informatics (ICBHI'17) scientific challenge respiratory sound database [44]. The dataset contains 920 recordings from 126 patients, and a total of 6,898 respiratory cycles: 1,864 cycles are annotated by respiratory experts as crackles, 886 cycles are identified as wheezes, 506 cycles are both, and the rest are normal. According to ICBHI official standards, 60% of breathing cycles are marked as the training set, and 40% make up the test set. All samples are recorded with different equipment from hospitals in Portugal and Greece by two different research teams. Most audio samples of the database were acquired by the research team of the Respiratory Research and Rehabilitation Laboratory (Lab3R) of the School of Health Sciences, University of Aveiro (ESSUA), and the others were recorded by the research team of the Aristotle University of Thessaloniki (AUTH). A significant



**TABLE 4.** Comparisons among different respiratory classification systems with ICBHI Challenge's official split (Highest scores in bold).

System	Test Cond.	Specificity(%)	Sensitivity(%)	Score(%)
GMM-HMM [12]	original split (60/40)	—	—	39.56%
Decision Tree [13]	original split (60/40)	75%	12%	43%
CNN-MoE [19]	original split (60/40)	68%	26%	47%
VGG-16(two-path) [16]	original split (60/40)	<b>81%</b>	28%	<b>54%</b>
ResNet	original split (60/40)	55%	<b>39%</b>	47%
SE-ResNet	original split (60/40)	69%	25%	47%
FB-SE-ResNet	original split (60/40)	61%	36%	48%
FBQ-SE-ResNet	original split (60/40)	70%	30%	50%

**FIGURE 12.** Examples of original FBank features (left: normal, center: crackle, right: wheeze).**FIGURE 13.** Examples of scaled FBank features with FB-attention (left: normal, center: crackle, right: wheeze).**FIGURE 14.** Examples of scaled FBank features with FBQ-attention (left: normal, center: crackle, right: wheeze).

Res layer. FB-SE-ResNet applies the FB attention module on the SE-ResNet framework. FBQ-SE-ResNet scales the FB attention module of FB-SE-ResNet according to the simulation results of bronchial modelling. We compared them with public systems [12], [13], [16], [18]–[20] in terms of the official evaluation metrics, Specificity, Sensitivity, and Score, which were released by the organization.

The details are shown in Table 4. Since the results in some references were calculated to two decimal places and the other results were whole integer values, we unified the accuracy to whole integer values.

### 1) SE BLOCK

From Table 4, the ResNet-backbone system gained 9% relative improvement in score value compared to the Decision Tree method [13], which is the baseline for the ICBHI 2017 Challenge. Based on the ResNet-backbone system, we introduced SE blocks and obtained relatively better performance in which the specificity value increased by 25%. The SE block's characteristics render it more attuned to feature channels between Res layers, thereby contributing to such improvements. However, the sensitivity value decreased by 35%, which may be due to unbalanced sample numbers

among the four classes; for example, the normal class contains as many samples as the total number from the other three classes.

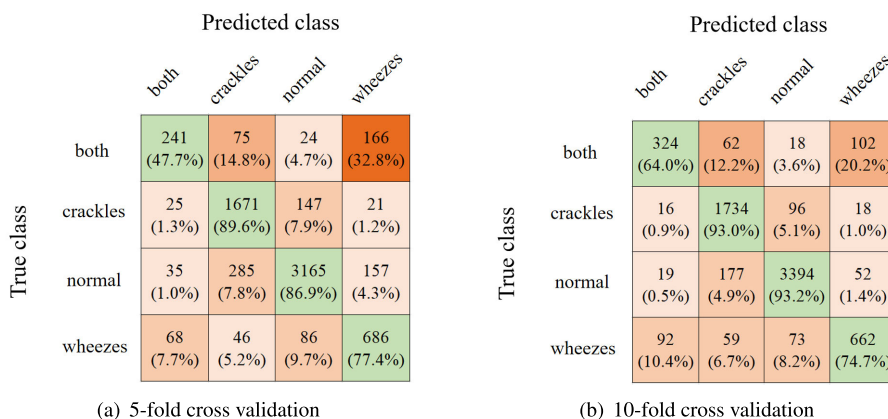
### 2) FEATURE-BAND ATTENTION MODULE

We focus on the effects of the feature-band attention module. In terms of the score value, the FB-SE-ResNet and FBQ-SE-ResNet systems were both better than the SE-ResNet system, which reached 2% and 6% improvements, respectively. The sensitivity values of the FB-SE-ResNet and FBQ-SE-ResNet systems were inferior to the ResNet-backbone but were beyond the SE-ResNet. For specificity values, the reverse applied. We think this is due to the effects of the feature-band attention module, which extracts innate feature representations for abnormal respiratory sounds even with an imbalanced sample size. In addition, we analysed the classification performance on different values of  $q$  in the range of 1.1 to 2.0, as shown in Fig. 11. We found that the  $q$  value of the feature had a great influence on the performance improvement and chose the best system with  $q = 1.3$  to compare with the FB-SE-ResNet system.

Examples of FBank acoustic features, the scaled FBank features with FB attention learning, and the scaled FBank

**TABLE 5. Comparisons between the proposed system with the state-of-the-art systems with the random data split(highest scores in bold).**

System	Test Cond.	Specificity(%)	Sensitivity(%)	Score(%)
CNN-MoE [19]	5-fold	90%	68%	79%
CNN-RNN [18]	10-fold	84.14%	48.63%	66.38%
CNN(parallel-pooling) [20]	10-fold	86%	61%	73.5%
FBQ-SE-ResNet	5-fold	87%	80%	83%
FBQ-SE-ResNet	10-fold	<b>93%</b>	<b>84%</b>	<b>88%</b>



**FIGURE 15. Confusion matrices for the FBQ-SE-ResNet.**

features with FBQ attention learning, which are extracted from the normal, crackle, and wheeze respiratory sounds respectively, are shown in Fig. 12 – 14. The original FBank features of abnormal lung sounds, such as wheezes and crackles, show significant features. And the scaled FBank features with the introduction of the feature-band attention module, allow the systems to focus on important acoustic feature bands, thereby improving the classification performance.

Table 4 also shows the performance obtained by the proposed FBQ-SE-ResNet framework and state-of-the-art published systems (where available), CNN-MoE and two-path VGG-16. We note that the FBQ-SE-ResNet framework lies second in the official training-set split, which only uses a low-complexity network structure ResNet-9, but is very competitive to the state-of-the-art systems.

For the experiment with 5-fold cross validation, we obtained a sensitivity value of 80%, a specificity value of 87%, and a score value of 83%. Compared to the CNN-MoE system, the proposed FBQ-SE-ResNet gained 18% relative improvement in sensitivity and 5% relative improvement in score, while there was a relative decrease in specificity of 3%. For the experiment with 10-fold cross validation, we achieved an outstanding performance with a sensitivity value of 93%, a specificity value of 84%, and a score value of 88%. The classification results are presented in the confusion matrices as shown in Fig. 15.

Let’s take for example, the 10-fold cross validation. It can be observed that the true positive rates of normal, crackles, wheezes and both (crackles and wheezes) are 93.0%, 93.0%, 74.7% and 64.0% respectively. In the case of the normal class,

4.9%, 1.4% and 0.5% of samples are falsely predicted as crackles, wheezes and both (crackles and wheezes) respectively. In the crackles class, 5.1%, 1.0% and 0.9% of samples are wrongly predicted as normal, wheezes and both (crackles and wheezes) respectively. In the wheezes class, 6.7%, 8.2% and 10.4% of samples are incorrectly identified as crackles, normal and both (crackles and wheeze) respectively. For the class of both (crackles and wheezes), only 3.6% of samples are falsely identified as normal, while 12.2% and 20.2% of samples are incorrectly classified as crackles and wheezes respectively. Although, it’s still an interesting and challenging task to further improve the discriminant performance of the class of both (crackles and wheezes), the results demonstrate the proposed system is a progressive and innovative approach to the diagnosis of normal and abnormal respiratory sounds.

**VI. CONCLUSION**

We studied the physical modelling of lung bronchi to simulate sound pressure curves of normal breath sounds and wheezes. The simulation results show that respiratory sound signals in the bronchi are mainly distributed below 2,000 Hz, and the spectrum peaks differ among various types of respiratory sounds. Based on such respiratory sound characteristics, we designed a feature-band attention module to adaptively weight the spectrum characteristics of the respiratory sound signals; this feature-band attention module is then used as the input of the end-to-end respiratory sound classification system. Experimental results on public databases indicate that the proposed end-to-end respiratory sound classification systems with a feature-band attention module achieve promising

performance. In the future, we intend to explore the impact of actual physiological conditions, such as increased airway mucus secretion and airway mucosal oedema. In addition, we plan to optimize the sensitivity value with data augmentation and the transfer learning technique.

## ACKNOWLEDGMENT

(Fuchuan Tong and Lingling Liu contributed equally to this work.)

## REFERENCES

- [1] E. Ray, D. Culliford, H. Kruk, K. Gillett, M. North, C. M. Astles, A. Hicks, M. Johnson, S. X. Lin, R. Orlando, M. Thomas, R. E. Jordan, D. Price, M. Konstantin, and T. M. A. Wilkinson, "Specialist respiratory outreach: A case-finding initiative for identifying undiagnosed COPD in primary care," *NPJ Primary Care Respiratory Med.*, vol. 31, no. 1, pp. 1–8, Dec. 2021.
- [2] L. Fraiwan, O. Hassanin, M. Fraiwan, B. Khassawneh, A. M. Ibnian, and M. Alkhodari, "Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers," *Biocybern. Biomed. Eng.*, vol. 41, no. 1, pp. 1–14, Jan. 2021.
- [3] J. Zhang, W. Ser, J. Yu, and T. T. Zhang, "A novel wheeze detection method for wearable monitoring systems," in *Proc. Int. Symp. Intell. Ubiquitous Comput. Educ.*, May 2009, pp. 331–334.
- [4] R. Liu, S. Cai, K. Zhang, and N. Hu, "Detection of adventitious respiratory sounds based on convolutional neural network," in *Proc. Int. Conf. Intell. Informat. Biomed. Sci. (ICIBMS)*, Nov. 2019, pp. 298–303.
- [5] B. E. D. Kingsbury, N. Morgan, and S. Greenberg, "Robust speech recognition using the modulation spectrogram," *Speech Commun.*, vol. 25, no. 1, pp. 117–132, 1998.
- [6] C. N. Campubí, F. J. H. Pericás, and M. G. Moreno, "On the decorrelation of filter-bank energies in speech recognition," in *Proc. EUROSPEECH 4th Eur. Conf. Speech Commun. Technol.*, Sep. 1995, pp. 1381–1384.
- [7] A. Rizal, R. Hidayat, and H. A. Nugroho, "Lung sound classification using Hjorth descriptor measurement on wavelet sub-bands," *J. Inf. Process. Syst.*, vol. 15, no. 5, pp. 1068–1081, 2019.
- [8] H. Hermansky, "Perceptual linear predictive (PLP) analysis of speech," *J. Acoust. Soc. Amer.*, vol. 87, no. 4, pp. 1738–1752, 1990.
- [9] N. Sengupta, M. Sahidullah, and G. Saha, "Lung sound classification using cepstral-based statistical features," *Comput. Biol. Med.*, vol. 75, pp. 118–129, Aug. 2016.
- [10] Y. P. Kahya, E. C. Guler, and S. Sahin, "Respiratory disease diagnosis using lung sounds," in *Proc. 19th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Magnificent Milestones Emerg. Opportunities Med. Eng.*, Nov. 1997, pp. 2051–2053.
- [11] I. Sen, M. Saraclar, and Y. P. Kahya, "A comparison of SVM and GMM-based classifier configurations for diagnostic classification of pulmonary sounds," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 7, pp. 1768–1776, Jul. 2015.
- [12] N. Jakovljević and T. Lončar-Turukalo, "Hidden Markov model based respiratory sound classification," in *Proc. Int. Conf. Biomed. Health Informat.* Singapore: Springer, 2017, pp. 39–43.
- [13] B. M. Rocha, D. Filos, L. Mendes, G. Serbes, S. Ulukaya, Y. P. Kahya, N. Jakovljevic, T. L. Turukalo, I. M. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, N. Maglaveras, R. P. Paiva, I. Chouvarda, and P. de Carvalho, "An open access database for the evaluation of respiratory sound classification algorithms," *Physiol. Meas.*, vol. 40, no. 3, Mar. 2019, Art. no. 035001.
- [14] L. Li, W. Xu, Q. Hong, F. Tong, and J. Wu, "Classification between normal and adventitious lung sounds using deep neural network," in *Proc. 10th Int. Symp. Chin. Spoken Lang. Process. (ISCSLP)*, Oct. 2016, pp. 1–5.
- [15] F. Demir, A. Sengur, and V. Bajaj, "Convolutional neural networks based efficient approach for classification of lung diseases," *Health Inf. Sci. Syst.*, vol. 8, no. 1, pp. 1–8, Dec. 2020.
- [16] K. Minami, H. Lu, H. Kim, S. Mabu, Y. Hirano, and S. Kido, "Automatic classification of large-scale respiratory sound dataset based on convolutional neural network," in *Proc. 19th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2019, pp. 804–807.
- [17] K. Minami, H. Lu, T. Kamiya, S. Mabu, and S. Kido, "Automatic classification of respiratory sounds based on convolutional neural network with multi images," in *Proc. 5th Int. Conf. Biomed. Imag., Signal Process.*, Sep. 2020, pp. 17–21.
- [18] J. Acharya and A. Basu, "Deep neural network for respiratory sound classification in wearable devices enabled by patient specific model tuning," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 3, pp. 535–544, Jun. 2020.
- [19] L. Pham, H. Phan, R. Palaniappan, A. Mertins, and I. McLoughlin, "CNN-MoE based framework for classification of respiratory anomalies and lung disease detection," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 8, pp. 2938–2947, Aug. 2021.
- [20] F. Demir, A. M. Ismael, and A. Sengur, "Classification of lung sounds with CNN model using parallel pooling structure," *IEEE Access*, vol. 8, pp. 105376–105383, 2020.
- [21] F. R. Haselton and P. W. Scherer, "Flow visualization of steady streaming in oscillatory flow through a bifurcating tube," *J. Fluid Mech.*, vol. 123, pp. 315–333, Oct. 1982.
- [22] B. Snyder and D. E. Olson, "Flow development in a model airway bronchus," *J. Fluid Mech.*, vol. 207, pp. 379–392, Oct. 1989.
- [23] Y.-S. Cheng, Y. Zhou, and B. T. Chen, "Particle deposition in a cast of human oral airways," *Aerosol Sci. Technol.*, vol. 31, no. 4, pp. 286–300, Oct. 1999.
- [24] L. Euler, "Principes généraux du mouvement des fluides," in *Mémoires de l'Académie des Sciences de Berlin*, vol. 1757, no. 226. Euler Archive All Works, pp. 274–315. [Online]. Available: <https://scholarlycommons.pacific.edu/euler-works/226>
- [25] A. Chorin, "Numerical solution of the Navier–Stokes equations," *Math. Comput.*, vol. 22, no. 104, pp. 745–762, 1968.
- [26] X. Zhao, X. Xu, D. Sun, S. Tan, F. Niu, Y. Liu, and X. Ren, "Numerical simulation for airflow structure and aerosol deposition in human upper respiratory," *J. Syst. Simul.*, vol. 24, no. 8, pp. 1582–1587, 2012.
- [27] X. Y. Luo, J. S. Hinton, T. T. Liew, and K. K. Tan, "LES modelling of flow in a simple airway model," *Med. Eng. Phys.*, vol. 26, no. 5, pp. 403–413, Jun. 2004.
- [28] M. Mihaescu, G. Mylavarapu, E. J. Gutmark, and N. B. Powell, "Large eddy simulation of the pharyngeal airflow associated with obstructive sleep apnea syndrome at pre and post-surgical treatment," *J. Biomech.*, vol. 44, no. 12, pp. 2221–2228, Aug. 2011.
- [29] J. C. Hardin and J. L. Patterson, "Monitoring the state of the human airways by analysis of respiratory sound," *Acta Astronautica*, vol. 6, no. 9, pp. 1137–1151, Sep. 1979.
- [30] C. A. Vaz and N. V. Thakor, "Adaptive Fourier estimation of time-varying evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 4, pp. 448–455, Apr. 1989.
- [31] J. Xu, Q. Chen, Y. Min, X. Zang, and B. Lu, "A study on wheezes models," *ACTA ACUSTICA, Chin. Version*, vol. 6, pp. 463–471, Dec. 1994.
- [32] E. Messner, M. Hagmuller, P. Swatek, F.-M. Smolle-Juttner, and F. Pernkopf, "Respiratory airflow estimation from lung sounds based on regression," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 1123–1127.
- [33] P. D. Muthusamy, K. Sundaraj, and N. A. Manap, "Computerized acoustical techniques for respiratory flow-sound analysis: A systematic review," *Artif. Intell. Rev.*, vol. 53, no. 5, pp. 3501–3574, Jun. 2020.
- [34] J. E. F. Williams and D. L. Hawkings, "Sound generation by turbulence and surfaces in arbitrary motion," *Philos. Trans. Roy. Soc. London A, Math. Phys. Sci.*, vol. 264, no. 1151, pp. 321–342, 1969.
- [35] Y. Zhang, H. Chen, K. Wang, and M. Wang, "Aeroacoustic prediction of a multi-element airfoil using wall-modeled large-eddy simulation," *AIAA J.*, vol. 55, no. 12, pp. 4219–4233, Dec. 2017.
- [36] X. Li, K. Yang, H. Hu, X. Wang, and S. Kang, "Effect of tailing-edge thickness on aerodynamic noise for wind turbine airfoil," *Energies*, vol. 12, no. 2, p. 270, Jan. 2019.
- [37] F. Shang, Z. Li, L. Li, and D. Guo, "Modeling of airflow-based lung sound for characterization of bronchi," in *Proc. 12th IEEE Int. Conf. Anti-Counterfeiting, Secur., Identificat. (ASID)*, Nov. 2018, pp. 245–249.
- [38] D. Sun, "Study on the vortex evolution and its relationship with the diffusion and transition of aerosol in human upper respiratory tract under the effect of fluid-solid interaction," Ph.D. dissertation, Acad. Military Med. Sci., Beijing, China, 2012.
- [39] D. Y. Ma, *Theoretical Basis of Modern Acoustics*. Beijing, China: Science Press, 2004, pp. 31–34.
- [40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [41] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

- [42] W. A. Wall and T. Rabczuk, "Fluid-structure interaction in lower airways of CT-based lung geometries," *Int. J. Numer. Methods Fluids*, vol. 57, no. 5, pp. 653-675, 2008.
- [43] X. Xu, D. Sun, F. Li, X. Zhao, and Y. Liu, "Numerical simulation on vortex evolution of airflow in realistic human upper respiratory tract under the effect of fluid-solid interaction," *Chin. J. Comput. Mech.*, vol. 32, no. 4, pp. 537-543, 2015.
- [44] B. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, and R. P. Paiva, "A respiratory sound database for the development of automated classification," in *Proc. Int. Conf. Biomed. Health Informat.*, Cham, Switzerland: Springer, 2017, pp. 33-37.



**QINGYANG HONG** (Member, IEEE) received the Ph.D. degree in computer science from the City University of Hong Kong, in January 2005. After that, he began his job at Xiamen University, where he is now an Associate Professor with the School of Informatics. He has published more than 60 papers and one book. His research interests include speaker recognition, speech recognition, and applications of deep learning.



**FUCHUAN TONG** received the B.S. degree in electronic and information engineering from Sichuan University, China, in 2019. He is currently pursuing the master's degree with the XMU Speech Laboratory, School of Electronic Science and Engineering, Xiamen University, China. His current research interests include acoustic signal processing, including speaker recognition and respiratory sound classification.



**LINGLING LIU** received the B.Eng. degree in electronic science and technology from Fuzhou University, China, in 2018, and the Master of Engineering degree from the School of Electronic Science and Engineering, Xiamen University, China. Her research interests include physical modeling and respiratory sound classification.



**XINGJIA XIE** received the B.S. degree in applied physics engineering from Fuzhou University, China, in 2021. He is currently pursuing the master's degree with the School of Electronic Science and Engineering, Xiamen University, China. His current research interests include deep learning and respiratory sound classification.



**LIN LI** (Member, IEEE) received the B.S. and Ph.D. degrees in electronic engineering from the University of Science and Technology of China, in 2003 and 2008, respectively. From 2008 to 2011, she was an Assistant Professor with the Electronic Engineering Department, Xiamen University, China. Since 2011, she has been an Associate Professor. She is currently the Vice Director with the Department of Electronic Engineering, Xiamen University. Her research interests include respiratory sound classification, speaker/language recognition, and speech signal processing and applications. She is a China Computer Federation (CCF) Member and a Committee Member of the National Conference on Man-Machine Speech Communication.

...